Phone vs. Face-to-Face with Virtual Persons

Timothy Bickmore
MIT Media Lab
20 Ames St., Room E15-320
Cambridge, MA 02139
+1 612 253 7368
bickmore@media.mit.edu

Justine Cassell
MIT Media Lab
20 Ames St., Room E15-315
Cambridge, MA 02139
+1 612 253 4899
justine@media.mit.edu

Abstract

This study compares people's interactions with Embodied Conversational Agents to similar interactions over the phone, and investigates the impact these media have on a wide range of behavioral, task and subjective measures. While the behavioral measures were consistent with previous studies, the subjective measures indicated that the fit of an ECA's persona to the task and style of interaction can overwhelm the effects of media on subjects' assessment of the ECA and the interaction.

Introduction

Social psychologists have compared face-to-face conversation with phone conversation, video-mediated communication and other mediated modalities, showing the effect various media have on psychosocial variables such as interpersonal vs. task-orientation, cooperation, trust, metacognition, person perception, veracity and task outcomes in negotiation and collaborative problem-solving (Rutter, 1987). Studies comparing human-human to human-computer interaction have demonstrated effects on speech disfluency, turn length and frequency, utterance length and interruptions (e.g., Oviatt, 1995). Few studies to date, however, have investigated how interaction with embodied conversational agents (ECAs) compares with these other well-understood modalities.

In this paper we present the results of a study comparing interaction with an embodied conversational agent to interaction with a phone-based dialogue system. This study extends previous work investigating the effects of social dialogue ("small talk") in a real estate sales domain, which demonstrated that social dialogue can have a significant impact on a user's trust of a computer agent (Bickmore and Cassell, 2001). In addition to varying medium (phone vs. embodied) and dialogue style (social dialogue vs. task-only) we also assessed the user's personality along the introversion/extroversion dimension, since extroversion is one indicator of a person's comfort level with face-to-face interaction.

1 Related Work

Work on the development of ECAs, as a distinct field of development, is best summarized in (Cassell, Sullivan et al., 2000). The current study is based on the REA ECA (see Figure 1), a simulated real-estate agent, who uses vision-based gesture recognition, speech recognition, discourse planning, sentence and gesture planning, speech synthesis and animation of a 3D body (Cassell, Bickmore et al., 1999). Some of the other major systems developed to date are Steve (Rickel and Johnson, 1998), the DFKI Persona (Andre, Muller et al., 1996), Olga (Beskow and McGlashan, 1997), and pedagogical agents developed by Lester, et al, (Lester, Stone et al., 1999). These systems vary in their linguistic generativity, input modalities, and task domains, but all aim to engage the user in natural, embodied conversation.

Figure 1. REA

1.2 User Studies on Embodied Conversational Agents

Koda and Maes (Koda and Maes, 1996) and Takeuchi and Naito (Takeuchi and Naito, 1995) studied interfaces with static or animated faces, and found that users rated them to be more engaging and entertaining than functionally equivalent interfaces without a face. Kiesler and Sproull (Kiesler and Sproull, 1997) found that users were more likely to be cooperative with an interface agent when it had a human face (vs. a dog or cartoon dog).

Andre, Rist and Muller found that users rated their animated presentation agent ("PPP Persona") as more entertaining and helpful than an equivalent interface without the agent (Andre, Rist et al., 1998). However, there was no difference in actual performance (comprehension and recall of presented material) in interfaces with the agent vs. interfaces without it.

In a user study of the Gandalf system (Cassell and Thorisson, 1999), users rated the smoothness of the interaction and the agent's language skills significantly higher under test conditions in which Gandalf utilized limited conversational behavior (gaze, turn-taking and beat gesture) than when these behaviors were disabled.

Sproull et al. (Sproull, Subramani et al., 1997) showed that subjects rated a female embodied interface significantly lower in sociability and gave it a significantly more negative social evaluation compared to a text-only interface. Subjects also reported themselves to be more aroused (less relaxed and assured) when interacting with the embodied interface than when interacting with the text interface. They also presented themselves in a more positive light (gave themselves significantly higher scores on social desirability scales) and disclosed less (wrote significantly less and skipped more questions in response to queries by the interface) when interacting with an embodied interface vs. a text-only interface. Men were found to disclose more in the embodied condition and women disclosed more in the text-only condition.

Most of these evaluations have tried to address whether embodiment of a system is useful at all, by including or not including an animated figure. In their survey of user studies on embodied agents, Dehn and van Mulken conclude that there is no "persona effect", that is a general advantage of an interface with an animated agent over one without an animated agent (Dehn and Mulken, 1999). However, they believe that lack of evidence and inconsistencies in the studies performed to date may be attributable to methodological shortcomings and variations in the kinds of animations used, the kinds of comparisons made (control conditions), the specific measures used for the dependent variables, and the task and context of the interaction.

1.3 User Studies on Human-Human vs. Human-Computer Communication

Several studies have shown that people speak differently to a computer than another person, even though there are typically no differences in task outcomes in these evaluations. Hauptmann and Rudnicky (Hauptmann and Rudnicky, 1988) performed one of the first studies in this area. They asked subjects to carry out a simple information-gathering task through a (simulated) natural language speech interface, and compared this with speech to a co-present human in the same task. They found that speech to the simulated computer system was telegraphic and formal, approximating a command language. In particular, when speaking to what they believed to be a computer, subject's utterances used a small vocabulary, often sounding like system commands, with very few task-unrelated utterances, and fewer filled pauses and other disfluencies.

These results were extended in research conducted by Oviatt (Oviatt, 1995; Oviatt, Levow et al., 1998; Oviatt and Cohen, 2000), in which she found that speech to a computer system was characterized by a low rate of disfluencies relative to speech to a co-present human. She also noted that visual feedback has an effect on disfluency: telephone calls have a higher rate of disfluency than co-present dialogue. From these results, it seems that people speak more carefully and less naturally when interacting with a computer.

Boyle and Anderson (Boyle, Anderson et al., 1994) compared pairs of subjects working on a map-based task who were visible to each other with pairs of subjects who were co-present but could not see each other. Although no performance difference was found between the two conditions, when subjects could not see one another, they compensated by giving more verbal feedback and using longer utterances. Their conversation was found to be less smooth than that between mutually visible partners, indicated by more interruptions, and less efficient, as more turns were required to complete the task. The researchers concluded that visual feedback improves the smoothness and efficiency of the interaction, but that we have devices to compensate for this when visibility is restricted.

Daly-Jones, et al. (Daly-Jones, Monk et al., 1998), also failed to find any difference in performance between video-mediated and audio-mediated conversations, although they did find differences in the quality of the interactions (e.g., more explicit questions in audio-only condition).

Whittaker and O'Conaill (Whittaker and O'Conaill, 1997) survey the results of several studies which compared video-mediated communication with audio-only communication and concluded that the visual channel does not significantly impact performance outcomes in task-oriented collaborations, although it does affect social and affective dimensions of communication. Comparing video-mediated communication to face-to-face and audio-only conversations, they also found that speakers used more formal turn-taking techniques in the video condition even though users reported that they perceived many benefits to video conferencing relative to the audio-only mode.

1.4 Trait-based Variation in User Responses

Several studies have shown that users react differently to social agents based on their own personality and other dispositional traits. For example, Reeves and Nass have shown that users like agents that match their own personality (on the introversion/ extraversion dimension) more than those which do not, regardless of whether the personality is portrayed through text or speech (Reeves and Nass, 1996; Nass and Lee, 2000). Resnick and Lammers showed that in order to change user behavior via corrective error messages, the messages should have different degrees of "humanness" depending on whether the user has high or low self-esteem ("computer-ese" messages should be used with low self-esteem users, while "human-like" messages should be used with high-esteem users) (Resnick and Lammers, 1985). Rickenberg and Reeves showed that different types of animated agents affected the anxiety level of users differentially as a function of whether users tended towards internal or external locus of control [20].

2. Experimental Methods

This was a multivariate, multiple-factor, between-subjects experimental design, involving 58 subjects (69% male and 31% female).

2.1 Apparatus

One wall of the experiment room was a rear-projection screen. In the EMBODIED condition Rea appeared life-sized on the screen, in front of the 3D virtual apartments she showed, and her synthetic voice was played through two speakers on the floor in front of the screen. In the PHONE condition only the 3D virtual apartments were displayed and subjects interacted with Rea over an ordinary telephone placed on a table in front of the screen.

For the purpose of this experiment, Rea was controlled via a wizard-of-oz setup on another computer positioned behind the projection screen. The interaction script included verbal and nonverbal behavior specifications for Rea (e.g., gesture and gaze commands as well as speech), and embedded commands describing when different rooms in the virtual apartments should be shown. Three pieces of information obtained from the user during the interview were entered into the control system by the wizard: the city the subject wanted to live in; the number of bedrooms s/he wanted; and how much s/he was willing to spend. The first apartment shown was in the specified city, but had twice as many bedrooms as the subject requested and cost twice as much as s/he could afford (they were also told the price was "firm"). The second apartment shown was in the specified city, had the exact number of bedrooms requested, but cost 50% more than the subject could afford (but this time, the subject was told that the price was "negotiable"). The scripts for the TASK and SOCIAL conditions were identical, except that the SOCIAL script had additional small talk utterances added to it, as described in (Bickmore and Cassell, 2001). The part of the script governing the dialogue from the showing of the second apartment through the end of the interaction was identical in both conditions.

Procedure. Subjects were told that they would be interacting with Rea, who played the role of a real

estate agent and could show them apartments she had for rent. They were told that they were to play the role of someone looking for an apartment in the Boston area. In both conditions subjects were told that they could talk to Rea "just like you would to another person".

2.2 Measures

Subjective evaluations of Rea -- including how friendly, credible, lifelike, warm, competent, reliable, efficient, informed, knowledgeable and intelligent she was -- were measured by single items on nine-point Likert scales. Evaluations of the interaction--including how tedious, involving, enjoyable, natural, satisfying, fun, engaging, comfortable and successful it was--were also measured on nine-point Likert scales. Evaluation of how well subjects felt they knew Rea, how well she knew and understood them and how close they felt to her were measured in the same manner.

Liking of REA was an index composed of three items--how likeable and pleasant Rea was and how much subjects liked her--measured items on nine-point Likert scales (Cronbach's alpha = .87).

Amount Willing to Pay was computed as follows. During the interview, Rea asked subjects how much they were able to pay for an apartment; subjects' responses were entered as $X per month. Rea then offered the second apartment for $Y (where Y = 1.5 X), and mentioned that the price was negotiable. On the questionnaire, subjects were asked how much they would be willing to pay for the second apartment, and this was encoded as Z. The task measure used was (Z - X) / (Y - X), which varies from 0% if the user did not budge from their original requested price, to 100% if they offered the full asking price.

Trust was measured by a standardized trust scale (Wheeless and Grotz, 1977) (alpha = .93).

Given literature on the relationship between user personality and preference for computer behavior, we were concerned that subjects might respond differentially based on predisposition. Thus, we also included composite measures for introversion and extroversion on the questionnaire.

Extrovertedness was an index composed of seven Wiggins (Wiggins, 1979) extrovert adjective items: Cheerful, Enthusiastic, Extroverted, Jovial, Outgoing, and Perky. It was used for assessment of the subject's personality (alpha = .87).

Introvertedness was an index composed of seven Wiggins (Wiggins, 1979) introvert adjective items: Bashful, Introverted, Inward, Shy, Undemonstrative, Unrevealing, and Unsparkling. It was used for assessment of the subject's personality (alpha = .84).

Behavioral Measures

Rates of speech disfluency (as defined in Oviatt, 1995) and utterance length were coded from the video data.

Observation of the videotaped data made it clear that some subjects took the initiative in the conversation, while others allowed Rea to lead. Unfortunately, Rea is not yet able to deal with user-initiated talk, and so user initiative often led to Rea interrupting the speaker. To assess the effect of this phenomenon, we therefore divided subjects into *PASSIVE* (below the mean on number of user-initiated utterances) and *ACTIVE* (above the mean on number of user-initiated utterances). To our surprise, these measures turned out to be independent of introversion/extroversion (Pearson r=0.042), and to not be predicted by these latter variables.

3. Results

Full factorial single measure ANOVAs were run, with SOCIALITY (Task vs. Social), PERSONALITY OF SUBJECT (Introvert vs. Extrovert), MEDIUM (Phone vs. Embodied) and INITIATION (Active vs. Passive) as independent variables.

3.1. Subjective Assessments of Rea

In looking at the questionnaire data we find that subjects seemed to feel more comfortable interacting with Rea over the phone than face-to-face. Thus, subjects in the phone condition felt that they knew Rea better ($F=5.02$; $p<.05$), liked her more ($F=4.70$; $p<.05$), felt closer to her ($F=13.37$; $p<.001$), felt more comfortable with the interaction ($F=3.59$; $p<.07$), and thought Rea was more friendly ($F=8.65$;$p<.005$), warm ($F=6.72$; $p<.05$), informed ($F=5.73$; $p<.05$), and knowledgeable ($F=3.86$; $p<.06$) than those in the embodied condition.

However, in the remainder of the results section, as we look more closely at different users, different kinds of dialogue styles, and users' actual behaviour, a more complicated picture emerges. Subjects felt that Rea knew them ($F=3.95$; $p<.06$) and understood them ($F=7.13$; $p<.05$) better when she used task-only dialogue face-to-face; these trends were reversed for phone-based interactions. Task-only dialogue was more fun ($F=3.36$; $p<.08$) and less tedious ($F=8.77$; $p<.005$; see Figure 2) when embodied, while social dialogue was more fun and less tedious on the phone.

That is, in the face-to-face condition, subjects preferred Rea to simply "get down to business."

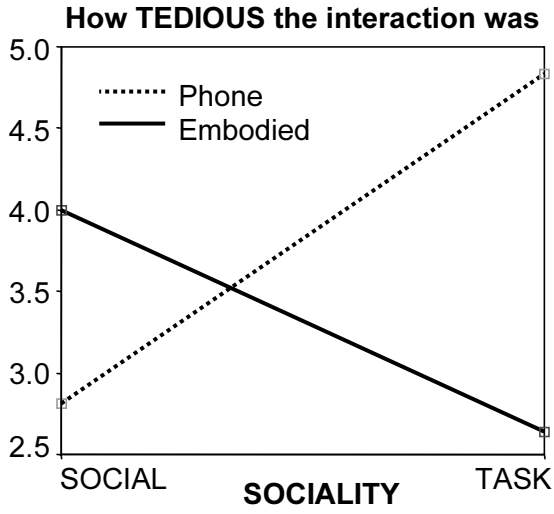

Figure 2. Ratings of TEDIOUS

These results may be telling us that Rea's nonverbal behavior inadvertently projected an unfriendly, introverted personality that was especially inappropriate for social dialogue. Rea's smiles are limited to those related to the ends of turns, and at the time of this experiment, she did not have a model of immediacy or other nonverbal cues for liking and warmth typical of social interaction (Argyle, 1988). According to Whittaker and O'Connail (Whittaker and O'Conaill, 1993), nonverbal information is especially crucial in interactions involving affective cues, such as negotiation or relational dialogue, and less important in purely problem-solving tasks. This interpretation of the results is backed up by comments such as this response from a subject in the face-to-face social condition:

> The only problem was how she would respond. She would pause then just say "OK", or "Yes". Also when she looked to the side and then back before saying something was a little bit unnatural.

This may explain why subjects preferred task interactions face-to-face, while on the phone Rea's social dialogue had its intended effect of making subjects feel that they knew REA better, that she understood them better, and that the experience was more fun and less tedious.

In our earlier study, looking only at an embodied interface, we reported that extroverts trusted the system more when it engaged in small talk, while introverts were not affected by the use of small talk (Bickmore and Cassell, 2001). In the current study, these results were re-confirmed, but only in the embodied interaction; that is, a three-way interaction between SOCIALITY, PERSONALITY and MEDIUM ($F=3.96$; $p<.06$) indicated that extroverts trusted Rea more when she used social dialogue in embodied interactions, but there was essentially no effect of user's personality and social dialogue on trust in phone interactions. Further analysis of the data indicated that this result derived from the substantial difference between introverts and extroverts in the face-to-face task-only condition. Introverts trusted her significantly more in the face-to-face task-only condition than in the other conditions ($p<.03$), while extroverts trusted her significantly less in this condition than in the other conditions ($p.<01$).

In light of these new observations, our earlier results indicating that social dialogue leads to increased trust (for extroverts at least) needs to be revised. This further analysis indicates that the effects we observed may be due to the attraction of a computer displaying similar personality characteristics, rather than the process of trust-building. In the face-to-face, task-only condition both verbal and nonverbal channels were clearly indicating that Rea was an introvert (also supported by the comments that REA's gaze-away behavior was too frequent, an indication of introversion (Wilson, 1977)), and in this condition we find the introverts trusting more, and extroverts trusting less. In all other conditions, the personality cues are either conflicting (a mismatch between verbal and nonverbal behavior has been demonstrated to be disconcerting to users (Nass, Isbister et al., 2000)) or only one channel of cues is available (i.e. on the phone), yielding trust ratings that are close to the overall mean.

There was, nevertheless, a preference by extroverts for social dialogue as demonstrated by the fact that, overall, extroverts liked Rea more when she used social dialogue, while introverts liked her more when she only talked about the task ($F=8.09$; $p<.01$).

Passive subjects felt more comfortable interacting with Rea than active subjects did, regardless of whether the interaction was face-to-face or on the phone, or whether Rea used social dialogue or not. Passive subjects said that they enjoyed the interaction more ($F=4.47$; $p<.05$), felt it was more successful ($F=6.04$; $p<.05$) and liked Rea more ($F=3.24$; $p<.08$), and that Rea was more intelligent ($F=3.40$; $p<.08$),

and knew them better (F=3.42; p<.08) than active subjects. These differences may be explained by the fixed-initiative dialogue model used in the WOZ script. Rea's interaction was designed for passive users--there was very little capability in the interaction script to respond to unanticipated user questions or statements--and user initiation attempts were typically met with uncooperative system responses or interruptions. But, given the choice between phone and face-to-face, passive users preferred to interact with Rea face-to-face: they rated her as more friendly (F=3.56; p<.07) and informed (F=6.30; p<.05) in this condition. Passive users also found the phone to be more tedious, while active users also found the phone to be less tedious (F=5.15; p<.05). Active users may have found the face-to-face condition particularly frustrating since processing delays may have led to the perception that the floor was open (inviting an initiation attempt), when in fact the wizard had already instructed Rea to produce her next utterance.

However, when interacting on the phone, active users differed from passive users in that active users felt she was more reliable when using social dialogue and passive users felt she was more reliable when using task-only dialogue. When interacting face-to-face with Rea, there was no such distinction between active and passive users (F=4.67; p<.05).

3.2. Effects on Task Measure

One of the most tantalizing results obtained is that extroverts were willing to pay more for the same apartment in the embodied condition, while introverts were willing to pay more over the phone (F=3.41; p<.08), as shown in Figure 3.

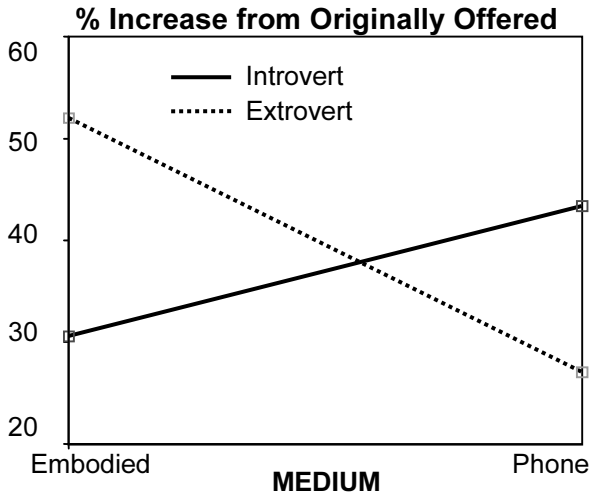

Figure 3. Amount Subjects Were Willing to Pay

While potentially very significant, this finding is a little difficult to explain, especially given that trust did not seem to play a role in the evaluation. Perhaps, since we asked our subjects to simply play the role of someone looking for an apartment, and given that the apartments displayed were cartoon renditions, the subjects may not have felt personally invested in the outcome, and thus may have been more likely to be persuaded by associative factors like the perceived liking and credibility of Rea. In fact, trust has been shown to not play a role in persuasion when "peripheral route" decisions are made, which is the case when the outcome is not of personal significance (Petty and Wegener, 1998). Further, extroverts are not only more sociable, but more impulsive than introverts (Wilson, 1977), and impulse buying is governed primarily by novelty (Onkvisit and Shaw, 1994). Extroverts did rate face-to-face interaction as more engaging than phone-based interaction (though not at a level of statistical significance), while introverts rated phone-based interactions as more engaging, providing some support for this explanation. It is also possible that this measure tells us more about subjects' assessment of the house than of the realtor. In future experiments we may ask more directly whether the subject felt that the realtor was asking a fair price.

3.3. Gender Effects

Women felt that Rea was more efficient (F=5.61; p<.05) and reliable (F=4.99; p<.05) in the embodied condition than when interacting with her over the phone, while men felt that she was more efficient and reliable by phone. Of course, Rea has a female body and a female voice and so in order to have a clearer picture of the meaning of these results, a similar study would need to be carried out with a male realtor.

3.4. Effects on Behavioral Measures

Although subjects' beliefs about Rea and about the interaction are important, it is at least equally important to look at how subjects *act*, independent of their conscious beliefs.

In this context we examined subjects' disfluencies when speaking with Rea. Remember that disfluency can be a measure of naturalness – human-human conversation demonstrates *more* disfluency than does human-computer communication. The rates of speech disfluencies (per 100 words) are shown in Table 1. Comparing these to results from previous studies (see

Table 2) indicates that interactions with REA were more similar to human-human conversation than to human-computer interaction. When asked if he was interacting with a computer or a person, one subject replied "A computer-person I guess. It was a lot like a human."

	Embodied	Phone	Overall
Disfluencies	4.83	6.73	5.83

Table 1. Speech Disfluencies per 100 Words

Human-human speech	
Two-person telephone call	8.83
Two-person face-to-face dialogue	5.5
Human-computer speech	
Unconstrained computer interaction	1.80
Structured computer interaction	0.83

Table 2. Speech Disfluencies per 100 Words for Different Types of Human-Human and Simulated Human-Computer Interactions (adapted from (Oviatt, 1995))

There were no significant differences in utterance length (MLU) across any of the conditions.
The behavioral measures indicate that, with respect to speech disfluency rates, talking to REA is more like talking to a person than talking to a computer.

Once again, there were significant effects of MEDIA, SOCIALITY and PERSONALITY on disfluency rate ($F=7.09$; $p<.05$), such that disfluency rates were higher in TASK than SOCIAL, higher overall for INTROVERTs than EXTROVERTs, higher for EXTROVERTs on the PHONE, and higher for INTROVERTs in EMBODIED condition. These effects on disfluency rates are consistent with the secondary hypothesis that the primary driver on disfluency is cognitive load, once the length of the utterance is controlled for (Oviatt, 1995). Given our results, this hypothesis would indicate that social dialogue requires lower cognitive load than task-oriented dialogue, that conversation requires a higher cognitive load on introverts than extraverts, that talking on the phone is more demanding than talking face-to-face for extraverts, and that talking face-to-face is more demanding than talking on the phone for introverts, all of which seem reasonable.

4. Conclusion

The complex results of this study give us hope for the future of embodied conversational agents, but also a clear roadmap for future research. In terms of their behaviour with Rea, users demonstrated that they treat conversation with her more like human-human conversation than like human-computer conversation. Their verbal disfluencies are the mark of unplanned speech, of a conversational style. However, in terms of their assessment of her abilities, this did not mean that users saw Rea through rose-colored glasses. They were clear about the necessity not only to embody the interaction, but to design every aspect of the embodiment in the service of the same interaction. That is, face-to-face conversations with ECAs must demonstrate the same quick timing of nonverbal behaviors as humans (not an easy task, given the state of the technologies we employ). In addition, the persona and nonverbal behavior of an ECA must be carefully designed to match the task, a conversational style, and user expectations. And finally, as computers begin to resemble humans, the bar of user expectations is raised: people expect that Rea will hold up her end of the conversation, including dealing with interruptions by active users.
We have begun to demonstrate the feasibility of embodied interfaces. Now it is time to design ECAs that people wish to spend time with, and that are able to use their bodies for conversational tasks for which human face-to-face interaction is unparalleled, such as social dialogue, initial business meetings, and negotiation.

Acknowledgements

Thanks to Ian Gouldstone, Jennifer Smith and Elisabeth Sylvan for help in conducting the experiment and analyzing data, and to the rest of the Gesture and Narrative Language Group for their help and support.

References

Andre, E., J. Muller, et al. (1996). The PPP Persona: A Multipurpose Animated Presentation Agent. Advanced Visual Interfaces, Palermo, Italy, ACM Press.

Andre, E., T. Rist, et al. (1998). Integrating Reactive and Scripted Behaviors in a Life-Like Presentation Agent. AGENTS '98, Minneapolis, MN.

Argyle, M. (1988). Bodily Communication. New York, Methuen & Co. Ltd.

Beskow, J. and S. McGlashan (1997). Olga: A Conversational Agent with Gestures. IJCAI'97 workshop on Animated Interface Agents - Making them Intelligent, Nagoya, Japan.

Bickmore, T. W. and J. Cassell (2001). Relational Agents: A Model and Implementation of Building User Trust. CHI 2001, Seattle, WA.

Boyle, E., A. Anderson, et al. (1994). "The Effects of Visibility in a Cooperative Problem Solving Task." Language and Speech **37**(1): 1-20.

Cassell, J., T. Bickmore, et al. (1999). Embodiment in Conversational Interfaces: Rea. CHI 99, Pittsburgh, PA, ACM.

Cassell, J., J. Sullivan, et al. (2000). Embodied Conversational Agents. Cambridge, MIT Press.

Cassell, J. and K. R. Thorisson (1999). "The Power of a Nod and a Glance: Envelope vs. Emotional Feedback in Animated Conversational Agents." Applied Artificial Intelligence **13**: 519-538.

Daly-Jones, O., A. Monk, et al. (1998). "Some advantages of video conferencing over high-quality audio conferencing: fluency and awareness of attentional focus." International Journal of Human-Computer Studies **49**(1): 21-58.

Dehn, D. M. and S. v. Mulken (1999). The Impact of Animated Interface Agents: A Review of Empirical Research. Saarbrucken, Germany, University of Saarland.

Hauptmann, A. G. and A. I. Rudnicky (1988). "Talking to computers: an empirical investigation." International Journal of Man-Machine Studies, **8**(6): 583 - 604.

Kiesler, S. and L. Sproull (1997). 'Social' Human-Computer Interaction. Human Values and the Design of Computer Technology. B. Friedman. Stanford, CA, CSLI Publications**:** 191-199.

Koda, T. and P. Maes (1996). Agents with Faces: The Effects of Personification of Agents. IEEE Robot-Human Communication '96, Tsukuba, Japan.

Lester, J. C., B. Stone, et al. (1999). "Lifelike Pedagogical agents for Mixed-Initiative Problem Solving in Constuctivist Learning Environments." User Modeling and User-Adapted Interaction **9**(1-2): 1-44.

Nass, C., K. Isbister, et al. (2000). Truth is Beauty: Researching Embodied Conversational Agents. Embodied Conversational Agents. J. Cassell, J. Sullivan, S. Prevost and E. Churchill. Cambridge, MA, MIT Press**:** 374-402.

Nass, C. and K. Lee (2000). Does Computer-Generated Speech Manifest Personality? An Experimental Test of Similarity-Attraction. CHI 2000, The Hague, Amsterdam, ACM Press.

Onkvisit, S. and J. J. Shaw (1994). Consumer Behavior: Strategy and Analysis. New York, Macmillan College Publishing Company.

Oviatt, S. (1995). "Predicting spoken disfluencies during human-computer interaction." Computer Speech and Language **9**: 19-35.

Oviatt, S. and P. Cohen (2000). "Multimodal Interfaces That Process What Comes Naturally." Communications of the ACM **43**(3): 45-53.

Oviatt, S., G.-A. Levow, et al. (1998). "Modeling global and focal hyperarticulation during human-computer error resolution." Journal of the Acoustical Society of America **104**(5): 3080-3091.

Petty, R. and D. Wegener (1998). Attitude Change: Multiple Roles for Persuasion Variables. The Handbook of Social Psychology. D. Gilbert, S. Fiske and G. Lindzey. New York, McGraw-Hill**:** 323-390.

Reeves, B. and C. Nass (1996). The Media Equation: how people treat computers, televisions and new media like real people and places. Cambridge, Cambridge University Press.

Resnick, P. V. and H. B. Lammers (1985). "The Influence of Self-esteem on Cognitive Responses to Machine-Like Versus Human-Like Computer Feedback." The Journal of Social Psychology **125**(6): 761--769.

Rickel, J. and W. L. Johnson (1998). "Task-Oriented Dialogs with Animated Agents in Virtual Reality." Proceedings of the 1st Workshop on Embodied Conversational Characters: 39-46.

Rickenberg, R. and B. Reeves (2000). The Effects of Animated Characters on Anxiety, Task Performance, and Evaluations of User Interfaces. CHI 2000, The Hague, Amsterdam.

Rutter, D. R. (1987). Communicating by Telephone. New York, Pergamon Press.

Sproull, L., M. Subramani, et al. (1997). When the Interface is a FAce. Human Values and the Design of Computer Technology. B. Friedman. Stanford, CA, CSLI Publications**:** 163-190.

Takeuchi, A. and T. Naito (1995). Situated Facial Displays: Towards Social Interaction. Human Factors in Computing Systems: CHI'95, Denver, CO, ACM Press.

Wheeless, L. and J. Grotz (1977). "The Measurement of Trust and Its Relationship to Self-Disclosure." Human Communication Research **3**(3): 250-257.

Whittaker, S. and B. O'Conaill (1993). An Evaluation of Video Mediated Communication. Conference on Human Factors and Computing Systems - INTERACT/CHI '93, Amsterdam, The Netherlands.

Whittaker, S. and B. O'Conaill (1997). The Role of Vision in Face-to-Face and Mediated Communication. Video-Mediated Communication. K. Finn, A. Sellen and S. Wilbur, Lawrence Erlbaum Associates, Inc.**:** 23-49.

Wiggins, J. (1979). "A psychological taxonomy of trait-descriptive terms." Journal of Personality and Social Psychology **37**(3): 395-412.

Wilson, G. (1977). Introversion/ Extraversion. Personality Variables in Social Behavior. T. Blass. New York, John Wiley & Sons**:** 179-218.